Fairness and Utilitarianism without Independence

Sinong Ma[1] [2]　　Zvi Safra[3] [4]

November 9, 2016

[1]The order of authors names was determined by a toss of a fair coin...

[2]University of Warwick; email: phd12sm@mail.wbs.ac.uk.

[3]University of Warwick (Emeritus, Tel Aviv University); email: zvi.safra@wbs.ac.uk.

[4]We are grateful to Simon Grant, Edi Karni, Tigran Melkonyan, Uzi Segal and Jean-Marc Tallon for very useful comments and suggestions.

Abstract

In this work we reconsider Harsanyi's celebrated (1953, 1955, 1977) utilitarian impartial observer theorem. Departing from Harsanyi's individualistic centered approach, we argue that, when societal decisions are at stake, postulates must not be drawn from individualistic behavior. Rather, they should be based on societal norms. Hence, notions like societal fairness should explicitly be the guiding principles. Continuing this line of thinking, we state and prove a utilitarian result that, rather than being based on the independence assumption, is based on the notion of procedural fairness and on similar treatment of societal and individual lotteries.

"An axiomatic justification of utilitarianism would have more content to it if it started off at a place somewhat more distinct from the ultimate destination" (Sen 1976, page 251)

1 Introduction

In this work we reconsider Harsanyi's celebrated (1953, 1955, 1977) utilitarian impartial observer theorem. We propose an approach that puts more emphasis on procedural fairness and we offer a utilitarian result that does not use the independence assumption.

Harsanyi analyzed a society that needs to choose among alternate social policies, each of which is a probability distribution over a given set of social actions, where the latter associate outcomes with the society's members. Every social lottery ℓ induces a lottery ℓ_i on individual i. Individual i's preferences $\succcurlyeq_i$ are known and different individuals may possess distinct preferences.

To help determine the optimal social policy, Harsanyi suggested that every individual is endowed with social preferences. Individuals may develop these preferences by adopting the role of an impartial observer, thus disregarding their true identities and acting behind "a veil of ignorance". Therefore, the impartial observer faces not only a lottery ℓ over social actions, but also a lottery γ over identities. By assumption, the impartial observer is able to compare situations in which two individuals get two different outcomes, under two disjoint social actions.

Harsanyi argued strongly for "Bayesian rationality". That is, he assumed that (among the other Bayesian postulates) all individuals satisfy the *independence assumption* of the expected utility theory, both at their personal and social preference layers. Harsanyi claimed that this "sound" axiom, together with the so-called *acceptance principle* (that an impartial observer fully adopts individual i's preferences if she imagines becoming that individual for sure), would force the impartial observer to be a (weighted) utilitarian. More formally, over all extended lotteries (γ, ℓ) in which the identity and the action lotteries are independently distributed, the impartial observer's preferences admit the following representation:

$$V(\gamma, \ell) = \sum_{i \in \mathcal{I}} \gamma_i U_i(\ell_i)$$

where γ_i is the probability of assuming person i's identity and $U_i(\ell_i) := \sum_x u_i(x)\ell_i(x)$ is person i's von Neumann-Morgenstern expected utility.

Like Harsanyi, most authors who derived modifications of the utilitarianism result within the impartial observer framework always assumed the independence axiom (see the works of Weymark (1991), Karni (1998) and Grant, Kajii, Polak and Safra (2010; henceforth GKPS)).[1] Notable exceptions within the related social aggregation framework are Blackorby, Donaldson and Mongin (2004) and Mongin and Pivato (2015).[2]

Interestingly, Harsanyi's entire emphasis on Bayesian rationality was based on an individual centered approach. Firstly, he assumed that rational individuals must satisfy the independence assumption and secondly, he claimed that society, by its need to be at least as rational as its members, must also satisfy independence (Harsanyi 1975). We disagree with Harsanyi on this. Instead we argue that when societal decision problems are at stake, postulates must not be drawn from individualistic behavior. Rather, they should be based on societal norms. Hence, when social preferences are formed, issues like societal fairness and equity should explicitly be the guiding principles.

In this work we focus on procedural fairness. This principle was first advocated by Diamond (1967) and was strongly supported by Sen (e.g., 1977). Its essence can be illustrated by the following example, which is an adoption of Diamond's example from the social aggregation framework to the impartial observer one. Consider a society that needs to decide on how to allocate an indivisible good between two individuals, 1 and 2, and let action a^i denotes allocating it to individual i. Suppose, as Diamond did, that $u_i(a^i) = 1$ for both i and $u_i(a^j) = 0$ for $i \neq j$ (that is, both individuals like the good, receive a utility of one unit from having it and zero otherwise). Also assume that the impartial observer considers

[1] A similar observation holds for most of the literature dealing with Harsanyi's social aggreagtion theorem. See Zhou (1997), Dhillon and Mertens (1999), Gilboa, Samet and Schmeidler (2004) and Fleurbaey and Mongin (2012).

[2] Unlike the other works (including the current one), these authors consider both *ex post* and *ex ante* analyses (and thus are able to employ Gorman's (1968) separability theorem).

equiprobable identity lotteries (that is, she imagines having equal chance of being either individual), evaluates all four outcomes in full agreement with the two individuals and adopts their utilities. The example can be described by the table

	a^1	a^2
1	1	0
2	0	1

where individuals 1 and 2 correspond to the rows, actions a^1 and a^2 correspond to the columns and the entries represent the impartial observer's utilities. The impartial observer has two policies at hand: Policy (1), which allocates the good to individual 1 (this policy is equivalent to choosing action a^1 and facing the first column of the table) and Policy (2), which allocates the good to one of the individuals, depending on the outcome of a toss of a fair coin (this policy is equivalent to the action lottery $\frac{1}{2}a^1 + \frac{1}{2}a^2$). The value of Policy (1) for Harsanyi's utilitarian observer is $\frac{1}{2} \times 1 + \frac{1}{2} \times 0 = \frac{1}{2}$, as is the value of Policy (2): $\frac{1}{2}\left(\frac{1}{2} \times 1 + \frac{1}{2} \times 0\right) + \frac{1}{2}\left(\frac{1}{2} \times 0 + \frac{1}{2} \times 1\right) = \frac{1}{2}$. Hence, the impartial observer is indifferent between the two policies.[3] However, Diamond and Sen argued that policy (2) provides both individuals with a "fair shake" and hence the impartial observer might prefer it.[4] This notion of procedural fairness is expressed in our work by the notion of *convexity* over action lotteries: if, given an identity lottery γ, two individuals disagree on the ranking of action lotteries ℓ and ℓ', then mixtures of these lotteries are weakly preferred over the less favorable one.[5]

Working in a framework in which the basic building blocks are two different types of lotteries, those over identities and those over actions, raises a natural question: should these separate types be treated similarly? Harsanyi, by applying the independence axiom to all possible pairs of identity-action lotteries, implicitly assumed that they should. Furthermore, in his own response to Diamond's concern about fairness, Harsanyi (1975) argued that even if randomizations were of value for promoting fairness (which he doubted), any explicit

[3] Note that the impartial observer is also indifferent between a^1 and a^2.

[4] A long list of real-life applications supporting Diamond's fairness consideration is provided by Elster (1989).

[5] Unlike Epstein and Segal (1992), we do not assume strict preference becuase, as was agrued by Sen (1977), mixture is not always superior.

randomization is superfluous since "the great lottery of (pre-)life" may be viewed as having already given each child an equal chance of being each individual. That is, it does not matter whether a good is allocated by a (possibly imaginary) lottery over identities or by a (real) lottery over actions. Put it differently, Harsanyi argues that we need to be indifferent between "accidents of birth" (identity lotteries) and real "life chances" (action lotteries). On this issue we follow the steps of Harsanyi and make this assumption explicit. We call it *source indifference.*

Despite its innocuous appearance, the conjunction of this assumption with procedural fairness turns out to be rather forceful. More precisely, the main result of this work shows that, assuming impartiality, convexity, source indifference and a stronger notion of acceptance are necessary, and sufficient, for utilitarianism.

Since the independence axiom is not assumed here, this result is novel and quite unexpected. Paraphrasing Sen's quote, we believe that one could hardly find an axiomatic justification of utilitarianism that starts off at a place that is more distinct from the ultimate destination than ours.

Lastly, our result implies that source indifference cannot hold if societies wish to exhibit *strict* inclination towards procedural fairness. Therefore, to accommodate views of authors like Diamond and Sen, the impartial observer must display preference for action lotteries over identity ones. We elaborate on this in the concluding section.

This work is organized as follows: Section 2 sets up the framework, Section 3 presents the assumptions, Section 4 states, and explains, the utilitarian result and section 5 concludes. Finally, proofs are given in Section 6.

2 Setup and Notation

Let $\mathcal{X} = [x_{\min}, x_{\max}] \subset \mathbb{R}$ be a compact interval representing all possible outcomes and let $\triangle(\mathcal{X})$ denote the set of outcome lotteries, endowed with the weak convergence topology. With slight abuse of notation, we will let x denote the degenerate outcome lottery that assigns probability 1 to outcome x. Let T be a denumerable set of potential individual types, where each type $t \in T$ is characterized by a preference relation $\succcurlyeq_t$ defined over

$\triangle(\mathcal{X})$. We assume throughout that each $\succcurlyeq_t$ is complete, transitive, continuous (in that the weak upper and lower contour sets are closed in the product topology), increases with respect to first-order stochastic-dominance and its asymmetric part $\succ_t$ is nonempty. The set of individuals under consideration is $\mathcal{I} = \cup_{t \in T} \mathcal{I}_t$, where $\mathcal{I}_t$ is a denumerable (infinite) set of type t individuals. A society I is a finite subset of $\mathcal{I}$. Note that, even though we allow for societies in which some individuals are of the same type, these individuals may receive different outcomes and hence they need not be treated similarly. Also note that our framework departs from Harsanyi's in that, instead of working with one fixed finite society, we consider all finite subsets of $\mathcal{I}$.[6]

A social policy, or an action, associates an outcome with every individual and hence is represented by a function $a : \mathcal{I} \to \mathcal{X}$. The set of all actions, endowed with the corresponding product topology, is denoted by $\mathcal{A}$ (two extreme actions, $a_{\max}$ and $a_{\min}$, defined by $a_{\max}(i) = x_{\max}$ and $a_{\min}(i) = x_{\min}$ for all i, respectively, will be used in the sequel). Let $\triangle(\mathcal{A})$ denote the set of simple lotteries (lotteries with finite support) over actions, with typical elements denoted by ℓ. With slight abuse of notation, we will let a denote the degenerate action lottery that assigns probability 1 to action a. A lottery $\ell \in \triangle(\mathcal{A})$ is sometimes written as $\ell = \sum_{a \in Supp(\ell)} \ell(a)\, a$.

Following Harsanyi, an observer imagines herself behind a veil of ignorance, uncertain about which identity she will assume in the given society. Let $\triangle(\mathcal{I})$ denote the set of simple identity lotteries on $\mathcal{I}$, where typical elements are denoted by γ (where γ_i is the probability assigned by the identity lottery γ to individual i). These lotteries represent the imaginary risks in the mind of the observer of being born as someone else. With slight abuse of notation, we will let i denote the degenerate identity lottery that assigns probability 1 to individual i. An imaginary lottery $\gamma \in \triangle(\mathcal{I})$ is sometimes written as $\gamma = \sum_{i \in Supp(\gamma)} \gamma_i i$. When the observer is faced with pairs of identity and action lotteries, it is assumed that they are independently distributed.

The observer is endowed with a preference relation $\succcurlyeq$ defined over the space of all product lotteries $\triangle(\mathcal{I}) \times \triangle(\mathcal{A})$. We assume throughout that $\succcurlyeq$ is complete, transitive, continuous and that its asymmetric part $\succ$ is nonempty. These assumptions imply that $\succcurlyeq$ admits

[6] The need for an infinite set of individuals is clarified in the proof of the theorem.

a (nontrivial) continuous representation $V : \triangle(\mathcal{I}) \times \triangle(\mathcal{A}) \rightarrow \mathbb{R}$. That is, for any pair of product lotteries (γ, ℓ) and (γ', ℓ'), $(\gamma, \ell) \succcurlyeq (\gamma', \ell')$ if and only if $V(\gamma, \ell) \geqslant V(\gamma', \ell')$. Note that the observer might prefer situations in which she imagines herself being a less privileged individual and receiving some amount x, over situations in which she imagines herself becoming a more privileged individual and receiving a larger amount x' (this may happen, for example, if she values affirmative action policies). As a result, there exists no *objective* natural order over the set of basic identity-outcome pairs (i, x) and, therefore, monotonicity with respect to first-order stochastic-dominance cannot be assumed. Instead, we require a weaker notion of monotonicity, based on the observer's *subjective* ranking over $\mathcal{I} \times \mathcal{X}$.

Definition 1. *Monotonicity*: For any pair of product lotteries (γ, a) and (γ', a'),

$$\sum_{\{i:V(i,a)\leqslant v\}} \gamma(i) \leqslant \sum_{\{i:V(i,a')\leqslant v\}} \gamma'(i) \quad \text{for all } v \in \operatorname{Im} V \quad \Rightarrow \quad (\gamma, a) \succcurlyeq (\gamma', a')$$

That is, a product lottery (γ, a) is preferred over another product lottery (γ', a') (both having degenerate action lotteries), if the probability of getting identity-action pairs with utilities not greater than v is always smaller under the first product.

For a given society I, let $\triangle(I)$ denote the set of identity lotteries over I.

Definition 2. *Utilitarianism*: The observer is a *utilitarian* if, for every society $I \subset \mathcal{I}$, her preferences restricted to $\triangle(I) \times \triangle(\mathcal{A})$ admit a representation of the form

$$V(\gamma, \ell) = \sum_{i\in I} \gamma_i U_i(\ell_i)$$

where $\ell_i \in \triangle(\mathcal{X})$ is the lottery faced by individual i (in which outcome $a(i)$ is assigned a probability of $\ell(a)$) and, for each individual i, $U_i(\ell_i) := \sum_{x\in\mathcal{X}} u_i(x)\,\ell_i(x)$ is an expected utility (EU) representation of $\succcurlyeq_i$.

As is well-known, the main behavioral property that characterizes EU preferences is *independence*:

Definition 3. *Independence*: Let $\succcurlyeq^*$ be a preference relation on $\triangle(\mathcal{X})$. Then, for all $p, q, r \in \triangle(\mathcal{X})$ and for all $\beta \in [0, 1]$,

$$p \succcurlyeq^* q \quad \Rightarrow \quad \beta p + (1-\beta) r \succcurlyeq^* \beta q + (1-\beta) r$$

3 Assumptions

We make the following assumptions on $\succcurlyeq$:

Axiom 1. *Impartiality*: For any two individuals $i, j \in \mathcal{I}$,

$$(1) \quad \text{for all} \quad \ell \in \triangle(\mathcal{A}), \quad \succcurlyeq_i = \succcurlyeq_j \quad \text{and} \quad \ell_i = \ell_j \quad \Rightarrow \quad (i, \ell) \sim (j, \ell)$$

$$(2) \quad (i, a_{\max}) \sim (j, a_{\max}) \quad \text{and} \quad (i, a_{\min}) \sim (j, a_{\min})$$

Part (1) of this axiom states that, given an action lottery ℓ, if two individuals i and j with identical preferences are faced with the same action lottery, then the observer is indifferent between facing ℓ, while being individual i, and facing ℓ, while being individual j. This requirement seems quite natural. Part (2) says that being individual i and getting the most preferred outcome $x_{\max}$ is assumed ethically equivalent to being individual j and getting the (same) most preferred outcome $x_{\max}$. As was convincingly explained by Karni (1998) who, in a different framework, employed a stronger axiom to derive utilitarianism "This value judgment ... is obtained by default. The methodological framework of revealed preference provides no ground for preferring one individual's most preferred alternative over that of the other. Consequently, strict preference in either direction is either biased or involves considerations other than the rank order of the alternatives". Clearly, the same applies to the worst outcome $x_{\min}$. A similar notion lies behind Segal's (2000) *dictatorship indifference* axiom.

Henceforth we assume that the observer preferences satisfy the impartiality axiom. To emphasize it, we call her an *impartial* observer.

Axiom 2. *Strong acceptance*: For all $i \in \mathcal{I}$ and $\ell, \ell' \in \Delta(\mathcal{A})$ satisfying $\forall j \neq i \quad \ell_j = \ell'_j$, if $\gamma_i > 0$ then

$$\ell_i \succcurlyeq_i \ell'_i \quad \Leftrightarrow \quad (\gamma, \ell) \succcurlyeq (\gamma, \ell')$$

This axiom states that the impartial observer sympathizes with individual i and fully adopts his preferences when she imagines herself being this individual with a positive probability, and when all other individuals are unaffected by her choice. This axiom strengthens Harsanyi's *acceptance* principle, according to which this sympathy holds for $\gamma_i = 1$. Axiom

2 also is equivalent to an axiom called *strong Pareto*, a version of Harsanyi's *Pareto* principle that was used in his aggregation analysis (see Harsanyi (1955), Weymark (1991) and Epstein and Segal (1992)).[7] To see the connection between our axiom and the strong Pareto principle note that, by sequentially applying our axiom, the following property holds: for *any* $\ell, \ell' \in \Delta(\mathcal{A})$, if $\ell_i \succcurlyeq_i \ell'_i$ for all $i \in Supp(\gamma)$ then $(\gamma, \ell) \succcurlyeq (\gamma, \ell')$.[8] In a sense, strong acceptance unifies two of Harsanyi's main ideas, taken from his two famous analyses of social choice theory. Finally, our axiom is analogous to Karni's (1998) *sympathy* assumption.

The strong acceptance axiom enables us to express the impartial observer's function V as a *social welfare function.* That is, V can be expressed as a function W that, instead of the action lottery ℓ, depends on the individuals' utilities associated with their induced lotteries ℓ_i. More formally, let $V_i(\ell_i) := V(i, \ell)$ be a representing utility the impartial observer attaches to individual i preferences. Note that, by impartiality, $V_i(x_{\min}) = V_j(x_{\min}) := v_{\min}$ and $V_i(x_{\max}) = V_j(x_{\max}) := v_{\max}$, for all $i, j \in \mathcal{I}$, and hence by continuity, the image of V_i, for all i, is equal to the closed interval $[v_{\min}, v_{\max}]$. Then, strong acceptance implies that $V(\gamma, \ell)$ can be written as $W\left(\vec{\gamma}, \vec{V}(\ell)\right)$, where W is defined over $\triangle([v_{\min}, v_{\max}])$, the set of lotteries over all attainable utility values in which, for all $i \in Supp(\gamma)$, $\vec{\gamma}_i = \gamma_i$ is the probability of attaining $\left(\vec{V}(\ell)\right)_i = V_i(\ell_i)$. To see how W is constructed assume, for expositional clarity, that $Supp(\gamma) = \{1, ..., n\}$. Then, given V and V_i, for any $\gamma \in \triangle(\{1, ..., n\})$ and $\vec{v} = (v_1, ..., v_n) \in [v_{\min}, v_{\max}]^n$, define W by $W(\vec{\gamma}, \vec{v}) := V(\gamma, \ell)$, for the imaginary lottery γ satisfying $\gamma_i = \vec{\gamma}_i$ and for any ℓ satisfying $v_i = V_i(\ell_i)$, for all $i \in \{1, ..., n\}$. By strong acceptance, W is well defined. Furthermore for a given $\vec{\gamma}$, W is monotonic increasing with respect to v_i whenever $\vec{\gamma}_i > 0$. Note that, by construction, W satisfies $W(1, v) = v$, for all $v \in [v_{\min}, v_{\max}]$.

The following properties will be used in the sequel.

[7] *Strong Pareto*: For a given society I, (1) for all lotteries $\ell, \ell' \in \Delta(\mathcal{A})$, if $\ell_i \succcurlyeq_i \ell'_i$ for all i, then $\ell \succcurlyeq \ell'$ and (2) if, furthermore, there exists an individual i' such that $\ell_{i'} \succ_{i'} \ell'_{i'}$, then $\ell \succ \ell'$.

[8] To see it, assume without loss of generality that $Supp(\gamma) = \{1, 2, ..., n\}$ and note that

$$(\gamma, \ell) = (\gamma, (\ell_1, \ell_2, ..., \ell_n)) \succcurlyeq (\gamma, (\ell'_1, \ell_2, ..., \ell_n)) \succcurlyeq (\gamma, (\ell'_1, \ell'_2, ..., \ell_n)) \succcurlyeq \cdots \succcurlyeq (\gamma, (\ell'_1, \ell'_2, ..., \ell'_n)) = (\gamma, \ell')$$

Lemma 1. Assume the observer satisfies *impartiality* and *strong acceptance.* Then
(a) for all $\ell, \ell' \in \triangle(\mathcal{A})$,

$$\succcurlyeq_i = \succcurlyeq_j \quad \text{and} \quad \ell_i = \ell'_j \quad \Rightarrow \quad (i, \ell) \sim (j, \ell')$$

(b) for all $(\gamma, \ell) \in \triangle(\mathcal{I}) \times \triangle(\mathcal{A})$,

$$(i, \ell) \sim (j, \ell) \quad \text{for all } i, j \in Supp(\gamma) \quad \Rightarrow \quad (\gamma, \ell) \sim (k, \ell), \quad \text{for all } k \in Supp(\gamma)$$

(c) for all $(\gamma^e, \ell), (\gamma^e, \ell') \in \triangle(\mathcal{I}) \times \triangle(\mathcal{A})$, where $Supp(\gamma^e) = \{1, ..., n\}$ and $\gamma^e = \sum_{i=1}^n \frac{1}{n} i$, if there exists a permutation π on $\{1, ..., n\}$ such that $(i, \ell_i) \sim \left(\pi(i), \ell'_{\pi(i)}\right)$ for all i, then

$$(\gamma^e, \ell) \sim (\gamma^e, \ell')$$

The proof appears in Section 6.1.

Axiom 3. *Convexity*: Consider two lotteries $\ell, \ell' \in \triangle(\mathcal{A})$ for which there exist two individuals i and j satisfying $\ell_i \succ_i \ell'_i$ and $\ell_j \prec_j \ell'_j$. Then, for any $\gamma \in \triangle(\mathcal{I})$ satisfying $\gamma_i > 0$, $\gamma_j > 0$ and for any $\beta \in (0, 1)$,

$$(\gamma, \ell) \succcurlyeq (\gamma, \ell') \quad \Rightarrow \quad (\gamma, \beta\ell + (1 - \beta)\ell') \succcurlyeq (\gamma, \ell')$$

That is, if two (participating) individuals disagree on the ranking of two action lotteries, then mixtures of these lotteries are weakly preferred over the less favorable one. As was explained in the introduction, this axiom is an expression of procedural fairness and is in agreement with Diamond's critique.

We include the requirement of having two individuals with opposing preferences since procedural fairness has greater appeal when real conflict exists. However, it is straightforward to verify that, with continuity, this requirement can be omitted. Hence, in situations where only one individual faces distinct lotteries under the action lotteries ℓ and ℓ', convexity implies that his preferences must also be convex.

Convexity is also related to social stability. Consider a society $I \subset \mathcal{I}$, whose set of available actions is given by a finite $A \subset \mathcal{A}$. For a given identity lottery $\gamma \in \triangle(I)$, the impartial observer's aim is to find the optimal action lottery that maximizes her utility. That is, the impartial observer seeks to solve the problem

$$\max_{\ell \in \triangle(A)} V(\gamma, \ell)$$

For societal stability, it is desirable that the set of optimal action lotteries does not change drastically when only minor changes occur. That is, we want this set to be upper hemi-continuous and convex valued with respect to the set of available actions A. Clearly, the continuity of $\succcurlyeq$ implies upper hemi-continuity, while convexity is equivalent to the optimal set being a convex valued correspondence.

Axiom 4. *Source indifference*: For all societies $\{i_1, ..., i_n\}$ and for all sets of available actions $\{a^1, ..., a^n\}$, if there exists $k \in \{1, ..., n\}$ such that $\left(i_j, a^k\right) \sim (i_k, a^j)$ for all j, then

$$\left(\gamma^e, a^k\right) \sim (i_k, \ell^e)$$

where $\gamma^e = \sum_{j=1}^{n} \frac{1}{n} i_j$ and $\ell^e = \sum_{j=1}^{n} \frac{1}{n} a^j$.

To illustrate, consider the following matrix

	a^1	a^2	$\cdots$	a^k	$\cdots$	a^n
i_1				x_1		
i_2				x_2		
$\vdots$				$\vdots$		
i_k	y^1	y^2	$\cdots$	z	$\cdots$	y^n
$\vdots$				$\vdots$		
i_n				x_n		

and suppose that the impartial observer is indifferent between the following two options, for all j: (1) receiving an outcome x_j while facing the deterministic action a^k and imagining being individual i_j, and (2) receiving an outcome y^j while facing the deterministic action a^j and imagining being individual i_k. There are two ways to randomize, with equal probabilities, over these degenerate pairs of equivalent product lotteries. The product lottery $\left(\gamma^e, a^k\right)$ randomizes over identity lotteries (for the given action a^k), while product lottery (i_k, ℓ^e) randomizes over action lotteries (for the given individual i_k). Then, as was argued by Harsanyi in his response to Diamond and was implicitly assumed by him, the impartial observer should be indifferent between the two randomizations.

The following lemma shows that, given impartiality and strong acceptance, source indifference for equi-probability lotteries γ^e and ℓ^e implies that this property holds for all lotteries γ and ℓ^γ. This property will be used later on.

Lemma 2. Assume the observer satisfies *impartiality*, *strong acceptance* and *source indifference*. For all societies $\{i_1, ..., i_n\}$ and for all sets of available actions $\{a^1, ..., a^n\}$, if there exists $k \in \{1, ..., n\}$ such that $\left(i_j, a^k\right) \sim (i_k, a^j)$ for all j, then, for all $\gamma = \sum_{j=1}^n \gamma_{i_j} i_j$ and $\ell^\gamma = \sum_{j=1}^n \gamma_{i_j} a^j$,

$$\left(\gamma, a^k\right) \sim (i_k, \ell^\gamma)$$

The proof is relegated to Section 6.1.

4 Utilitarianism

Our main result shows that the preceding axioms force all individuals to be of the EU type and, in addition, the impartial observer must be a utilitarian. That is, the behavioral assumptions on the impartial observer preferences induce her, as well as all individuals, to satisfy the independence axiom. This is achieved without imposing independence explicitly (neither on individuals nor on the observer).

Theorem. Assume the observer satisfies *impartiality*. Then her preferences satisfy *strong acceptance*, *convexity* and *source indifference* if, and only if, all individuals in $\mathcal{I}$ satisfy *independence* and the observer is a *utilitarian*.

The proof, which is relegated to Section 6.2, consists of two parts. First, we prove that all individuals in $\mathcal{I}$ must satisfy the independence axiom. Then, we demonstrate that the impartial observer's preferences can be represented by a function that is additive with respect to the identity lotteries and that she, too, must satisfy the independence axiom.

Comment 1. Consider the Diamond example, represented by the table

	a^1	a^2
1	1	0
2	0	1

Having the identity lottery $\gamma^e = \left(\frac{1}{2}, \frac{1}{2}\right)$, choosing action a^i corresponds to the pair (γ^e, a^i), while tossing a fair coin corresponds to the pair $(\gamma^e, \ell^e) = \left(\gamma^e, \frac{1}{2}a^1 + \frac{1}{2}a^2\right)$. By source indifference, $(\gamma^e, a^1) \sim (1, \ell^e)$ and $(\gamma^e, a^1) \sim (2, \ell^e)$. Hence, $(1, \ell^e) \sim (2, \ell^e)$ and therefore, by Lemma 1(b), $(1, \ell^e) \sim (\gamma^e, \ell^e)$. But then, by transitivity, $(\gamma^e, a^1) \sim (\gamma^e, \ell^e)$ and the impartial observer is indifferent between the first action (Policy

(1)) and the mixture (Policy (2)). Put differently, she does not strictly prefer tossing a fair coin over the pure action a^1. Moreover, it can now be seen (proof omitted) that, by convexity, *any* mixture of the two actions a^1 and a^2 must be indifferent to a^1. This may seem like a significant step towards proving utilitarianism. However, the derivation of these 'straight indifference line segments' from the above extremely symmetric situation does not extend to the general case and cannot be utilized to derive a utilitarian representation.

Comment 2. As noted in the introduction, Blackorby, Donaldson and Mongin (2004) and Mongin and Pivato (2015) also derived utilitarianism without imposing independence. Although these authors work within Harsanyi's aggregation theorem framework, a comparison to our theorem seems natural and is carried out by focusing on the analysis of Mongin and Pivato (2015). Consider a given society I, with a set of actions A, and identify every product lottery (γ, ℓ) with a matrix whose rows correspond to individuals and columns correspond to actions. Mongin and Pivato's *ex ante* analysis is manifested by their *row preference* assumption, an assumption that is analogous to our strong acceptance axiom. Similarly, their *ex post* analysis is manifested by a *column preference* assumption that, in our model, would require an improvement in the impartial observer situation whenever an action a is replaced by a better action $\bar{a}$. Together with a *coordinate monotonicity* assumption, these two assumptions enable Mongin and Pivato to employ Gorman's (1968) separability theorem and to derive a fully separable representation of the observer preferences. As can be seen in Section 6.2, our proof uses different arguments. Nevertheless, one might conjecture that, since source indifference implies similar treatment of columns and rows, then, together with strong acceptance, Gorman's separability theorem could be applied to yield our result. However, this is not true. As can be seen in Examples 1 and 2 below, strong acceptance and source indifference are not sufficient to imply utilitarianism.

Comment 3. Another result that is close to ours appears in GKPS (2010). Their Theorem 3 roughly states that an observer is a utilitarian if and only if she satisfies acceptance, independence over identity lotteries and (their notion) of source indifference. However, as we do not assume any form of independence, the current result is stronger than theirs.[9]

The following first two examples demonstrate the necessity of convexity. The third demonstrates the necessity of source indifference.

Example 1. Here we present a non utilitarian impartial observer who satisfies all axioms except for convexity. Assume that all preferences $\succcurlyeq_i$ of individuals $i \in \mathcal{I}$ belong to the rank-dependent utility class (RDU; see Weymark (1981) and Quiggin (1982)). Let $g : [0,1] \to [0,1]$ be an increasing and onto function. For a given simple lottery r and $z \in Supp\,(r)$ define $F_r\,(z) := \sum_{y \leqslant z} r\,(y)$, $F_r\,(z_{-}) := \sum_{y<z} r\,(y)$ and $\nabla g\,(z;r) := g\,(F_r\,(z)) - g\,(F_r\,(z_{-}))$. On simple lotteries, RDU preferences are represented by a function of the form $V\,(p) = \sum_x u\,(x)\,\nabla g\,(x;p)$. When g is the identity function, $\nabla g\,(x;p) = p\,(x)$ and RDU preferences are reduced to EU preferences. We assume that, in the eyes of the impartial observer, individual i's preferences are represented by $V_i\,(p) = \sum_x u_i\,(x)\,\nabla g\,(x;p)$, where g is common to all individuals and, for all $i,j \in \mathcal{I}$, $u_i\,(x_{\min}) = u_j\,(x_{\min})$ and $u_i\,(x_{\max}) = u_j\,(x_{\max})$. The observer preferences are also of the RDU type and are represented by

$$V^r\,(\gamma,\ell) = \sum_{i \in I} V_i\,(\ell_i)\,\nabla g\,(V_i\,(\ell_i)\,;\gamma)$$

Impartiality and strong acceptance are satisfied by construction. To verify that source indifference is satisfied consider, without loss of generality, a society $I = \{1,...,n\}$, a set of available actions $\{a^1,...,a^n\}$ and assume that there exists k for which $V^r\left(j,a^k\right) = V^r\left(k,a^j\right)$

[9]It should also be noted that the notion of source indifference used by GKPS (2010) (they termed it 'indifference between identity and action lotteries') is stronger than ours. This is formally stated as Lemma 4 (see Section 6.4).

for all j. Then, for all j,

$$u_j\left(a^k(j)\right)=V_j\left(a^k(j)\right)=V^r\left(j,a^k\right)=V^r\left(k,a^j\right)=V_k\left(a^j(k)\right)=u_k\left(a^j(k)\right)$$

Hence,

$$\begin{aligned}V^r\left(\gamma^e,a^k\right) &= \sum_{j\in I}u_j\left(a^k(j)\right)\nabla g\left(u_j\left(a^k(j)\right);\gamma^e\right)\\ &= \sum_{j\in I}u_k\left(a^j(k)\right)\nabla g\left(u_j\left(a^k(j)\right);\gamma^e\right)\\ &= \sum_{j\in I}u_k\left(a^j(k)\right)\nabla g\left(a^j(k);\ell^e_k\right)=V^r\left(k,\ell^e\right)\end{aligned}$$

as required.

To see that convexity does not hold assume that g is strictly concave and fix $j\in I$. Let $\ell,\ell'\in\Delta(\mathcal{A})$ be two distinct action lotteries satisfying $\ell_i=\ell'_i$ for all $i\neq j$, $\ell_j\neq\ell'_j$ and $V_j(\ell_j)=V_j\left(\ell'_j\right)$ (clearly, such lotteries exist). The strict concavity of g implies $V_j\left(\frac{1}{2}\ell_j+\frac{1}{2}\ell'_j\right)<V_j(\ell_j)$ and hence, for any γ with $\gamma_j>0$, $V^r\left(\gamma,\frac{1}{2}\ell+\frac{1}{2}\ell'\right)<V^r(\gamma,\ell)$.[10]

Note that, as the following case shows, non-convexity of $\succcurlyeq_i$ (which is manifested by the concavity of g), is not necessary for the non-convexity of $\succcurlyeq$. For this, let $I=\{1,...,5\}$, consider the two actions described by the following matrix (the entries are the utility values)

	a^1	a^2
1	1	0
2	0	1
3	1	1
4	1	1
5	1	1

let g be given by the convex piecewise linear function

$$g(t)=\begin{cases}0 & t\leqslant 0.2\\ -\frac{1}{4}+\frac{5}{4}t & \text{otherwise}\end{cases}$$

[10] Perhaps the simplest way to see the connection between strict concavity of g and quasi-convexity of V_j is to observe that, for continuous lotteries, $V_j(\ell_j)=u_j(x_{\max})-\int_z g\left(F_{\ell_j}(z)\right)u'_j(z)\,\mathrm{d}z$. This immediately implies $V_j\left(\frac{1}{2}\ell_j+\frac{1}{2}\ell'_j\right)<V_j(\ell_j)$. Similarly, if g is convex then so is $\succcurlyeq_j$.

and note that, by the convexity of g, each $\succcurlyeq_i$ is convex.

Clearly,

$$V^r\left(\gamma^e, a^j\right) = g\left(0.2\right) \times 0 + \left(1 - g\left(0.2\right)\right) \times 1 = 1$$

for both $j = 1, 2$. Next, consider the lottery $\frac{1}{2}a^1 + \frac{1}{2}a^2$. For $i \in \{1, 2\}$,

$$V_i\left(\frac{1}{2}a^1\left(i\right) + \frac{1}{2}a^2\left(i\right)\right) = g\left(0.5\right) \times 0 + \left(1 - g\left(0.5\right)\right) \times 1 = \frac{5}{8}$$

while, for $i \in \{3, 4, 5\}$, $V_i\left(\frac{1}{2}a^1\left(i\right) + \frac{1}{2}a^2\left(i\right)\right) = 1$. Hence, for the impartial observer,

$$\begin{aligned} V^r\left(\gamma^e, \frac{1}{2}a^1 + \frac{1}{2}a^2\right) &= g\left(0.4\right) \times \frac{5}{8} + \left(1 - g\left(0.4\right)\right) \times 1 \\ &= \frac{1}{4} \times \frac{5}{8} + \frac{3}{4} \times 1 = \frac{29}{32} < 1 \end{aligned}$$

and convexity is not satisfied.

Example 2. In the two cases described in Example 1, either individual preferences are non-convex with respect to outcome lotteries (when g is concave) or the impartial observer preferences are non-convex with respect to identity lotteries (when g is convex). This might suggest that convexity would be satisfied if all preferences involved were convex. As we now show, this conjecture is false.

Assume that individual preferences are weighted utility (WU; see Chew (1983)). That is, for all i and $p \in \triangle(\mathcal{X})$,

$$V_i\left(p\right) = V\left(p\right) = \sum_k p_k \frac{w\left(x_k\right)}{\sum_j p_j w\left(x_j\right)} u\left(x_k\right)$$

where u is a strictly increasing utility function and w is a non constant and positive weighting function. These preferences belong to the betweenness class (see Chew (1989) and Dekel (1986)), a class that is characterized by the property: for all lotteries p and q, $p \succcurlyeq q$ if and only if $p \succcurlyeq \lambda p + (1 - \lambda) q \succcurlyeq q$, for all $\lambda \in (0, 1)$. Clearly, betweenness implies that WU preferences are convex. Note that, although individuals have identical preferences over $\triangle(\mathcal{X})$, they do not necessarily agree on the ranking of action lotteries in $\triangle(\mathcal{A})$.

The impartial observer preferences are of the same type and are given by

$$V^w\left(\gamma, \ell\right) = \sum_i \gamma_i \frac{w\left(u^{-1}\left(V\left(\ell_i\right)\right)\right)}{\sum_j \gamma_j w\left(u^{-1}\left(V\left(\ell_j\right)\right)\right)} V\left(\ell_i\right)$$

As in Example 1, source indifference is satisfied. To see it, assume (for $k=1$) $V^w(j,a^1(j)) = V^w(1,a^j(1))$, for all j. That is, $u(a^1(j)) = u(a^j(1))$ or, equivalently, $a^1(j) = a^j(1)$, for all j.

Then

$$\begin{aligned} V^w\left(\gamma^e,a^1\right) &= \sum_i \frac{1}{n}\frac{w\left(\left(u^{-1}\circ u\right)\left(a^1(i)\right)\right)}{\sum_j \frac{1}{n} w\left(\left(u^{-1}\circ u\right)\left(a^1(j)\right)\right)} u\left(a^1(i)\right) \\ &= \sum_i \frac{1}{n}\frac{w\left(a^1(i)\right)}{\sum_j \frac{1}{n} w\left(a^1(j)\right)} u\left(a^1(i)\right) \\ &= \sum_i \frac{1}{n}\frac{w\left(a^i(1)\right)}{\sum_j \frac{1}{n} w\left(a^j(1)\right)} u\left(a^i(1)\right) \\ &= V^w\left(1,\ell^e\right) \end{aligned}$$

Next we show that convexity is not satisfied. Consider again the Diamond's example and assume that $w\left(u^{-1}(0)\right) = 0.1$, $w\left(u^{-1}\left(\frac{2}{3}\right)\right) = 0.6$, $w\left(u^{-1}\left(\frac{32}{33}\right)\right) = 0.75$ and $w\left(u^{-1}(1)\right) = 0.8$. Then,

$$V^w\left(\gamma^e,a^1\right) = \frac{\frac{1}{2}w\left(u^{-1}(1)\right)}{\frac{1}{2}w\left(u^{-1}(1)\right)+\frac{1}{2}w\left(u^{-1}(0)\right)} = \frac{0.5\times 0.8}{0.5\times 0.8+0.5\times 0.1} = \frac{8}{9}$$

and

$$V^w\left(\gamma^e,a^2\right) = \frac{\frac{1}{2}w\left(u^{-1}(1)\right)}{\frac{1}{2}w\left(u^{-1}(0)\right)+\frac{1}{2}w\left(u^{-1}(1)\right)} = \frac{0.5\times 0.8}{0.5\times 0.1+0.5\times 0.8} = \frac{8}{9}$$

Let $\ell = 0.8a^1 + 0.2a^2$ be a mixture of a^1 and a^2. Then,

$$\begin{aligned} V(\ell_1) &= \frac{0.8w\left(u^{-1}(1)\right)}{0.8w\left(u^{-1}(1)\right)+0.2w\left(u^{-1}(0)\right)} = \frac{0.8\times 0.8}{0.8\times 0.8+0.2\times 0.1} = \frac{32}{33} \\ V(\ell_2) &= \frac{0.2w\left(u^{-1}(1)\right)}{0.2w\left(u^{-1}(1)\right)+0.8w\left(u^{-1}(0)\right)} = \frac{0.2\times 0.8}{0.2\times 0.8+0.8\times 0.1} = \frac{2}{3} \end{aligned}$$

and, for the impartial observer,

$$\begin{aligned} V^w\left(\gamma^e,\ell\right) &= \frac{1}{2}\frac{w\left(u^{-1}\left(V(\ell_1)\right)\right)}{\frac{1}{2}w\left(u^{-1}\left(V(\ell_1)\right)\right)+\frac{1}{2}w\left(u^{-1}\left(V(\ell_2)\right)\right)}V(\ell_1) \\ &\quad +\frac{1}{2}\frac{w\left(u^{-1}\left(V(\ell_2)\right)\right)}{\frac{1}{2}w\left(u^{-1}\left(V(\ell_1)\right)\right)+\frac{1}{2}w\left(u^{-1}\left(V(\ell_2)\right)\right)}V(\ell_2) \\ &= \frac{0.75}{0.75+0.6}\times\frac{32}{33}+\frac{0.6}{0.75+0.6}\times\frac{2}{3} \\ &\approx 0.835 < \frac{8}{9} \end{aligned}$$

Hence, convexity is violated.

Example 3. A non utilitarian impartial observer who satisfies all axioms except for source indifference is the *generalized utilitarian* impartial observer of GKPS (2010). Consider

$$V^g(\gamma,\ell)=\sum_{i\in I}\gamma_i\phi_i\left[U_i(\ell_i)\right]$$

where $\phi_i:[v_{\min},v_{\max}]\to\mathbb{R}$ are strictly concave, for all i. It is easy to verify that strong acceptance and convexity are satisfied while, as was shown in GKPS, this observer deems identity lotteries inferior to action lotteries.

Comment 4. Consider the following assumption, which is weaker than source indifference.

Preference for identity lotteries: For all societies $\{i_1,...,i_n\}$ and for all sets of available actions $\{a^1,...,a^n\}$, if there exists $k\in\{1,...,n\}$ such that $\left(i_j,a^k\right)\sim\left(i_k,a^j\right)$ for all j, then

$$\left(\gamma^e,a^k\right)\succcurlyeq(i_k,\ell^e)$$

In Section 6.3 (Lemma 3) we show that this assumption, in conjunction with strong acceptance and convexity, implies source indifference. Therefore, our theorem could be stated in a slightly stronger form. The current statement is preferred because source indifference seems to be more natural than a preference for one type of lotteries over the other.

5 Conclusion

As stated in the introduction we argue that, when societal decisions are at stake, postulates must be drawn from society centered behavior. We have chosen to focus on the notion of procedural fairness (exhibited by convexity) and added to it the requirement that the impartial observer is indifferent between identity and action lotteries. In our main result we have shown that these two assumptions (together with strong acceptance) were sufficient to force the impartial observer to be a utilitarian. Unlike most utilitarian results, no form of the independence axiom was required here.

In addition to offering a society centered basis for utilitarianism, our result sheds more light on what is needed in order to always have a strict preference for procedural fairness.

Since preference for identity lotteries implies source indifference (Lemma 3, Section 6.2), then, in order to have a strict preference for procedural fairness, the impartial observer must display a preference for action lotteries. Two such non-utilitarian models exist in the literature. The first follows from Karni and Safra (2002).[11] In their model, which leads to the representation $V(\gamma, \ell) = \sum_{i \in I} \gamma_i V_i(\ell_i)$, individuals possess a sense of justice and the preference for procedural fairness is solely manifested by their behavior (their utilities V_i are assumed to be concave). It can easily be verified that this impartial observer displays a preference for action lotteries. The second model is the generalized utilitarian impartial observer of GKPS (2010). As mentioned above, GKPS show that a preference for action lotteries holds if and only if each ϕ_i is concave, a condition that implies procedural fairness. For a third model, consider a rank dependent, or a Gini, impartial observer, whose preferences are represented by

$$V^{rd}(\gamma, \ell) = \sum_{i \in I} \phi(U_i(\ell_i)) \nabla g(U_i(\ell_i); \gamma)$$

(where each U_i is of the EU type and both ϕ and g are concave). As can easily be verified, a preference for action lotteries follows from Chew, Karni and Safra (1987) while procedural fairness follows from Quiggin (1993, Section 9.1).

6 Proofs

6.1 Proofs of Lemmata 1 and 2

Proof of Lemma 1

(a) Assume $\succcurlyeq_i = \succcurlyeq_j$ and consider $\ell, \ell' \in \triangle(\mathcal{A})$ satisfying $\ell_i = \ell'_j$. Construct an action lottery $\bar{\ell}$ that satisfies $\bar{\ell}_i = \bar{\ell}_j = \ell_i = \ell'_j$. Then

$$(i, \ell) \sim \left(i, \bar{\ell}\right) \sim \left(j, \bar{\ell}\right) \sim (j, \ell')$$

as required (the first and the last indifferences follow from strong acceptance while the second follows from impartiality).

[11] See also Grant, Kajii, Polak and Safra (2012).

(b) Let $v = V(i, \ell) = V_i(\ell_i)$ and note that, by the arguments that precede the statement of the lemma, $V(\gamma, \ell) = W(\vec{\gamma}, (v, ..., v))$ while $V(k, \ell) = W(1, v)$. That is, the product lottery (γ, ℓ) is equivalent to a utility lottery with n identical outcomes (where n is the number of elements in $Supp(\gamma)$), all equal to v, while (k, ℓ) is equivalent to the degenerate lottery that yields v for sure. The two utility lotteries seem identical but, in order to show that the impartial observer is indeed indifferent between them, the monotonicity property must be employed.

For this, let $c_i(\ell_i) \in \mathcal{X}$ be individual i's certainty equivalent of the lottery ℓ_i (that is, $c_i(\ell_i) \sim_i \ell_i$) and consider the action $\hat{a}$ satisfying $\hat{a}(i) = c_i(\ell_i)$. By strong acceptance, $(\gamma, \ell) \sim (\gamma, \hat{a})$ and $(k, \ell) \sim (k, \hat{a})$. Then, as the unique utility value attained by both $(\gamma, \hat{a})$ and $(k, \hat{a})$ is v, monotonicity implies that $(\gamma, \hat{a}) \sim (k, \hat{a})$. By transitivity, $(\gamma, \ell) \sim (k, \ell)$.

(c) Let (γ^e, ℓ), (γ^e, ℓ') and π satisfy the conditions of the lemma. Construct two actions $\hat{a}$ and $\hat{a}'$ satisfying $\hat{a}(i) = c_i(\ell_i)$ and $\hat{a}'(i) = c_i(\ell'_i)$ where, as above, c_i is the certainty equivalent function of individual i. By strong acceptance, $(\gamma^e, \ell) \sim (\gamma^e, \hat{a})$ and $(\gamma^e, \ell') \sim (\gamma^e, \hat{a}')$. The conditions $(i, \ell_i) \sim \left(\pi(i), \ell'_{\pi(i)}\right)$ imply $V(i, \ell_i) = V\left(\pi(i), \ell'_{\pi(i)}\right)$ for all i, and hence,

$$V(i, \hat{a}) = V(i, c_i(\ell_i)) = V(i, \ell_i) = V\left(\pi(i), \ell'_{\pi(i)}\right) = V\left(\pi(i), c_{\pi(i)}\left(\ell'_{\pi(i)}\right)\right) = V(\pi(i), \hat{a}')$$

By monotonicity, $(\gamma^e, \hat{a}) \sim (\gamma^e, \hat{a}')$ and, by transitivity, $(\gamma^e, \ell) \sim (\gamma^e, \ell')$. ■

Proof of Lemma 2 Consider, without loss of generality, a society $I = \{1, ..., n\}$, a set of available actions $A = \{a^1, ..., a^n\}$ and assume that (again, without loss of generality) $(i, a^1) \sim (1, a^i)$, for all i. Let $\gamma = (\gamma_1, ..., \gamma_n)$.

First assume that γ is rational. That is, $\gamma_i = \frac{n_i}{m_i}$, for all i. Consider a new society $\bar{I} = \{\bar{1}, \bar{2}, ...\}$ with $m_1 \cdots m_n$ individuals, in which the first $n_1 m_2 \cdots m_n$ individuals are identical to individual 1 of I, the next $m_1 n_2 m_3 \cdots m_n$ individuals are identical to individual 2 of I, and so on. Similarly, let the set of actions $\bar{A} = \{\bar{a}^1, \bar{a}^2, ...\}$ consists of $m_1 \cdots m_n$ actions, in which the first $n_1 m_2 \cdots m_n$ actions are identical to action a^1 of A, the next $m_1 n_2 m_3 \cdots m_n$ actions are identical to action a^2 of A, and so on. Finally, let $\bar{\gamma}^e$ and $\bar{\ell}^e$ be the equi-probability lotteries over $\bar{I}$ and $\bar{A}$, respectively. By construction, $(\bar{\imath}, \bar{a}^1) \sim (\bar{1}, \bar{a}^{\bar{\imath}})$, for all $\bar{\imath}$. By source indifference, $(\bar{\gamma}^e, \bar{a}^1) \sim \left(\bar{1}, \bar{\ell}^e\right)$. To conclude note that, by monotonicity, $(\gamma, a^1) \sim (\bar{\gamma}^e, \bar{a}^1)$ and, by Lemma 1(a), $\left(\bar{1}, \bar{\ell}^e\right) \sim (1, \ell^\gamma)$. Transitivity then implies $(\gamma, a^1) \sim (1, \ell^\gamma)$.

Next consider any γ and let $\beta_k \rightarrow_{k\rightarrow\infty} \gamma$ be a sequence of rational lotteries that converge to γ. By construction, $(\beta_k, a^1) \rightarrow_{k\rightarrow\infty} (\gamma, a^1)$ and $\left(1, \ell^{\beta_k}\right) \rightarrow_{k\rightarrow\infty} (1, \ell^{\gamma})$. By the argument above, $(\beta_k, a^1) \sim \left(1, \ell^{\beta_k}\right)$ for all k and hence, by continuity, $(\gamma, a^1) \sim (1, \ell^{\gamma})$. ■

6.2 Proof of the Theorem

The 'if' part is immediate. The proof of the converse is divided into two parts.

Part I[12] In this part we show that all individuals satisfy the independence axiom. Consider an individual $i^* \in \mathcal{I}$ and denote his preferences by $\succcurlyeq^*$. We want to demonstrate that for all $p, q, r \in \triangle(\mathcal{X})$, $p \sim^* q \Rightarrow \frac{1}{2}p + \frac{1}{2}r \sim^* \frac{1}{2}q + \frac{1}{2}r$. This, using Herstein and Milnor (1953), would imply that $\succcurlyeq^*$ satisfies the independence axiom. Using the continuity of $\succcurlyeq^*$, we can restrict attention to equi-probability lotteries with the same number of outcomes: $p = \left(\left(\frac{1}{k}, ..., \frac{1}{k}\right), x\right)$, $q = \left(\left(\frac{1}{k}, ..., \frac{1}{k}\right), y\right)$, and $r = \left(\left(\frac{1}{k}, ..., \frac{1}{k}\right), z\right)$ (to see it, note that (1) any lottery with rational probabilities can be replicated by an equi-probability lottery with not necessarily distinct outcomes and (2) the set of lotteries with rational probabilities is dense in the space of all lotteries).

Consider a society I consisting of $n = 2k$ individuals, all with preferences $\succcurlyeq_i = \succcurlyeq^*$. Let $\pi_1 = (1, 2, ..., n)$, $\pi_2 = (2, 3, ..., 1)$,..., $\pi_n = (n, 1, 2, ..., n-1)$ be permutations on $\{1, ..., n\}$ (where $\pi_j(i)$ stands for the ith element of the permutation π_j). We concentrate on a set of actions $\dot{A} = \{\dot{a}^1, ..., \dot{a}^n\}$ available to the society that are defined as follows: for $j = 1, ..., k$

$$\dot{a}^j(i) = \begin{cases} x_{\pi_j(i)} & \text{if } 1 \leqslant i \leqslant k \\ z_{\pi_j(i-k)} & \text{if } k < i \leqslant n \end{cases}$$

and, for $j = k+1, ..., n$

$$\dot{a}^j(i) = \begin{cases} z_{\pi_{j-k}(i)} & \text{if } 1 \leqslant i \leqslant k \\ x_{\pi_{j-k}(i-k)} & \text{if } k < i \leqslant n \end{cases}$$

To illustrate, look at the following table

[12] The proof of this part is similar to that of Dekel, Safra and Segal (1991, Theorem 2). However, dealing with social multi-person frameowrk, our proof is more general than (and improves upon) theirs.

	$\dot{a}^1$	$\dot{a}^2$	$\cdots$	$\dot{a}^k$	$\dot{a}^{k+1}$	$\dot{a}^{k+2}$	$\cdots$	$\dot{a}^n$
1	x_1	x_2	$\cdots$	x_k	z_1	z_2	$\cdots$	z_k
2	x_2	x_3	$\cdots$	x_1	z_2	z_3	$\cdots$	z_1
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
k	x_k	x_1	$\cdots$	x_{k-1}	z_k	z_1	$\cdots$	z_{k-1}
$k+1$	z_1	z_2	$\cdots$	z_k	x_1	x_2	$\cdots$	x_k
$k+2$	z_2	z_3	$\cdots$	z_1	x_2	x_3	$\cdots$	x_1
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
n	z_k	z_1	$\cdots$	z_{k-1}	x_k	x_1	$\cdots$	x_{k-1}

Fact 1 $(\gamma^e, \ell^e) \sim (\gamma^e, \dot{a}^1)$.

Since for all i, j $\ell_i^e = \ell_j^e$, impartiality implies $(i, \ell^e) \sim (j, \ell^e)$ and hence, by Lemma 1(b), $(\gamma^e, \ell^e) \sim (1, \ell^e)$. Next, since $\dot{a}^j(1) = \dot{a}^1(j)$ (x_j if $j \leqslant k$ and z_{j-k} otherwise) then, in both $(1, \dot{a}^j)$ and $(j, \dot{a}^1)$, the impartial observer faces the same deterministic outcome. By Lemma 1(a), $(1, \dot{a}^j) \sim (j, \dot{a}^1)$ for all $j \in I$ and, by source indifference, $(1, \ell^e) \sim (\gamma^e, \dot{a}^1)$. Transitivity then implies $(\gamma^e, \ell^e) \sim (\gamma^e, \dot{a}^1)$.

Fact 2 Let $\ell^k = \frac{1}{k}\sum_{j=1}^{k} \dot{a}^j$. Then $\left(\gamma^e, \ell^k\right) \sim (\gamma^e, \ell^e)$.

Since all actions $\dot{a}^i$ yield the same outcomes then, using impartiality and monotonicity, $(\gamma^e, \dot{a}^i) \sim (\gamma^e, \dot{a}^1)$ for all i. By repeated application of convexity, $\left(\gamma^e, \ell^k\right) = \left(\gamma^e, \frac{1}{k}\sum_{j=1}^{k} \dot{a}^j\right) \succcurlyeq (\gamma^e, \dot{a}^1)$.[13] Hence, by Fact 1 and transitivity, $\left(\gamma^e, \ell^k\right) \succcurlyeq (\gamma^e, \ell^e)$.

For the converse, consider the action lottery $\hat{\ell}^k = \frac{1}{k}\sum_{j=k+1}^{n} \dot{a}^j$. For all $i = 1, ..., k$, $\hat{\ell}_i^k$, the lottery individual i faces under $\hat{\ell}^k$, is identical to ℓ_{k+i}^k, the lottery that individual $k+i$ faces under ℓ^k. By Lemma 1(a), $\left(i, \hat{\ell}_i^k\right) \sim \left(k+i, \ell_{k+i}^k\right)$. Similarly, $\hat{\ell}_{k+i}^k$, the lottery individual $k+i$ faces under $\hat{\ell}^k$, is identical to ℓ_i^k, the lottery that individual i faces under ℓ^k and hence, by Lemma 1(a), $\left(k+i, \hat{\ell}_{k+i}^k\right) \sim \left(i, \ell_i^k\right)$. Therefore, by Lemma 1(c), $\left(\gamma^e, \hat{\ell}^k\right) \sim \left(\gamma^e, \ell^k\right)$. Since $\ell^e = \frac{1}{2}\hat{\ell}^k + \frac{1}{2}\ell^k$, convexity implies $(\gamma^e, \ell^e) \succcurlyeq \left(\gamma^e, \ell^k\right)$.

Hence, $\left(\gamma^e, \ell^k\right) \sim (\gamma^e, \ell^e)$.

[13] Note that by continuity, the convexity axiom holds even when there are no opposing individuals (see Section 3, right after the statement of the convexity axiom).

Fact 3 $\frac{1}{2}p+\frac{1}{2}r \sim^* \frac{1}{2}q+\frac{1}{2}r$.

By the first part of the proof of Fact 1, $(\gamma^e,\ell^e) \sim (1,\ell^e)$. Therefore, using transitivity and Fact 2, $\left(\gamma^e,\ell^k\right) \sim (1,\ell^e)$. Note that in the first lottery, the first k individuals face the lottery p and the rest face the lottery r while, in the second, individual 1 is faced with the lottery $\frac{1}{2}p+\frac{1}{2}r$.

Next consider the same set of individuals I with another set of actions $\tilde{A}=\left\{\tilde{a}^1,...,\tilde{a}^{2k}\right\}$, that is derived from $\dot{A}$ by replacing every x_j by y_j. Clearly, a similar conclusion holds: the impartial observer is indifferent between the product lottery $\left(\gamma^e,\tilde{\ell}^k\right)$, in which the first k individuals face the lottery q and the rest face the lottery r, and the product lottery $\left(1,\tilde{\ell}^e\right)$, in which individual 1 is faced with the lottery $\frac{1}{2}q+\frac{1}{2}r$. But as $p \sim^* q$, all individuals in I are indifferent between p and q and hence, by strong acceptance, $\left(\gamma^e,\ell^k\right) \sim \left(\gamma^e,\tilde{\ell}^k\right)$. By transitivity, $(1,\ell^e) \sim \left(1,\tilde{\ell}^e\right)$. Hence the impartial observer, while imagining herself being individual 1, is indifferent between the lotteries $\frac{1}{2}p+\frac{1}{2}r$ and $\frac{1}{2}q+\frac{1}{2}r$. By strong acceptance, $\frac{1}{2}p+\frac{1}{2}r \sim^* \frac{1}{2}q+\frac{1}{2}r$.

To conclude Part I, note that allowing k to go to infinity implies that $\succcurlyeq^*$ satisfies independence over the entire set of lotteries $\triangle(\mathcal{X})$.[14]

Part II In the second part we show that the impartial observer is a utilitarian. Consider a society I (without loss of generality, $I=\{1,...,n\}$) and let $V(\gamma,\ell)$ be a representation of the impartial observer preferences where $\left(\vec{V}(\ell)\right)_i = V_i(\ell_i) = \varphi_i(U_i(\ell_i))$, φ_i is monotonic increasing and, by Part I, $U_i(\ell_i)=\sum_{x\in\mathcal{X}} u_i(x)\,\ell_i(x)$ is an EU representation of individual i's preferences. Since u_i is determined up to (positive) affine transformations, we can assume it satisfies $u_i(x_{\min})=v_{\min}$ and $u_i(x_{\max})=v_{\max}$ (hence, $\varphi_i(v_{\min})=v_{\min}$ and $\varphi_i(v_{\max})=v_{\max}$, for all i).

Fact 4 $\succcurlyeq$ can be represented by a separable function $\bar{V}(\gamma,\ell)=\sum_{i=1}^n \gamma_i\phi_i[U_i(\ell_i)]$.

Choose $(\gamma,\ell)\in\Delta(I)\times\Delta(\mathcal{A})$, denote $v_i=\varphi_i(U_i(\ell_i))$ and let $c_i(\ell_i)\in\mathcal{X}$ be individual i's certainty equivalent of the lottery ℓ_i (that is, $u_i(c_i(\ell_i))=U_i(\ell_i)$). Consider a set of actions $\hat{A}=\{\hat{a}^j \mid j\in\{1,...,n\}\}$ satisfying $\hat{a}^1(i)=c_i(\ell_i)$ and $\hat{a}^j(1)=(\varphi_1\circ u_1)^{-1}(v_j)$ for $i,j=1,...,n$. By construction, $V(i,\hat{a}^1)=(\varphi_i\circ u_i)(c_i(\ell_i))=v_i$ and $V(1,\hat{a}^i)=(\varphi_1\circ u_1)\circ$

[14] This is where we make use of the infinity of the set $\mathcal{I}$.

$(\varphi_1 \circ u_1)^{-1}(v_i) = v_i$. Hence $(i, \hat{a}^1) \sim (1, \hat{a}^i)$ and, by source indifference, $(\gamma, \hat{a}^1) \sim (1, \ell^\gamma)$ (ℓ^γ is the action lottery on $\hat{A}$ associated with γ). Put differently, $V(\gamma, \hat{a}^1) = V(1, \ell^\gamma)$. Note that by strong acceptance, $V(\gamma, \ell) = V(\gamma, \hat{a}^1)$. Therefore,

$$
\begin{aligned}
V(\gamma, \ell) &= V\left(\gamma, \hat{a}^1\right) = V(1, \ell^\gamma) = \varphi_1\left(U_1\left(\ell_1^\gamma\right)\right) \\
&= \varphi_1\left(\sum_{i=1}^n \gamma_i u_1\left((\varphi_1 \circ u_1)^{-1}(v_i)\right)\right) \\
&= \varphi_1\left(\sum_{i=1}^n \gamma_i \varphi_1^{-1}(v_i)\right) \\
&= \varphi_1\left(\sum_{i=1}^n \gamma_i \left(\varphi_1^{-1} \circ \varphi_i\right)\left(U_i\left(\ell_i\right)\right)\right)
\end{aligned}
$$

Denote $\bar{V} = \varphi_1^{-1} \circ V$ and $\phi_i = \varphi_1^{-1} \circ \varphi_i$ (note that $\bar{V}$ also represents the impartial observer preferences and its image is $[v_{\min}, v_{\max}]$). By the above,

$$
\bar{V}(\gamma, \ell) = \sum_{i=1}^n \gamma_i \phi_i\left[U_i\left(\ell_i\right)\right]
$$

Fact 5 $\succcurlyeq$ can be represented by the affine function $\bar{V}(\gamma, \ell) = \sum_{i=1}^n \gamma_i U_i(\ell_i)$.

To conclude, we show that for all i, $\bar{V}_i = \phi_i \circ U_i$ is affine which, given $\varphi_i(v_{\min}) = v_{\min}$ and $\varphi_i(v_{\max}) = v_{\max}$, implies $\bar{V}_i = U_i$. Take $\ell, \ell' \in \Delta(\mathcal{A})$. Since U_i is of the EU type, we have that for all $\lambda \in [0,1]$,

$$
\begin{aligned}
\bar{V}_i\left(\lambda \ell_i + (1-\lambda)\ell_i'\right) &= \phi_i\left[U_i\left(\lambda \ell_i + (1-\lambda)\ell_i'\right)\right] = \phi_i\left[\lambda U_i(\ell_i) + (1-\lambda) U_i(\ell_i')\right] \\
&= \phi_i\left[\lambda u_i(c_i(\ell_i)) + (1-\lambda) u_i(c_i(\ell_i'))\right] = \phi_i\left[U_i\left(\lambda c_i(\ell_i) + (1-\lambda) c_i(\ell_i')\right)\right] \\
&= \bar{V}_i\left(\lambda \breve{a}^i(i) + (1-\lambda)\breve{a}^j(i)\right) = \bar{V}\left(i, \lambda \breve{a}^i + (1-\lambda)\breve{a}^j\right)
\end{aligned}
$$

for actions $\breve{a}^i$ and $\breve{a}^j$ satisfying $\breve{a}^i(i) = c_i(\ell_i)$, $\breve{a}^j(i) = c_i(\ell_i')$ (note that the element $\lambda c_i(\ell_i) + (1-\lambda) c_i(\ell_i')$ that appears in the second line is a lottery, not an outcome). Defining $\breve{a}^i(j) = \left(\phi_j \circ u_j\right)^{-1} \circ (\phi_i \circ u_i)(c_i(\ell_i'))$ we get

$$
\bar{V}\left(j, \breve{a}^i\right) = \left(\phi_j \circ u_j\right) \circ \left(\phi_j \circ u_j\right)^{-1} \circ (\phi_i \circ u_i)(c_i(\ell_i')) = (\phi_i \circ u_i)(c_i(\ell_i')) = \bar{V}\left(i, \breve{a}^j\right)
$$

and hence, by source indifference and for γ satisfying $\gamma_i = \lambda$, $\gamma_j = 1-\lambda$ and $\gamma_k = 0$ otherwise,

$$
\bar{V}\left(i, \lambda \breve{a}^i + (1-\lambda)\breve{a}^j\right) = \bar{V}\left(\lambda i + (1-\lambda) j, \breve{a}^i\right)
$$

(note that actions $\breve{a}^k$ for $k \neq i, j$ are irrelevant but can easily be defined so as to fit with the requirements of the axiom). Now, by the structure of $\bar{V}$ and by using the equation $\bar{V}(j, \breve{a}^i) = \bar{V}(i, \breve{a}^j)$,

$$\begin{aligned}
\bar{V}\left(\lambda i + (1-\lambda) j, \breve{a}^i\right) &= \lambda \bar{V}\left(i, \breve{a}^i\right) + (1-\lambda) \bar{V}\left(j, \breve{a}^i\right) = \lambda \bar{V}\left(i, \breve{a}^i\right) + (1-\lambda) \bar{V}\left(i, \breve{a}^j\right) \\
&= \lambda \bar{V}_i\left(\breve{a}^i(i)\right) + (1-\lambda) \bar{V}_i\left(\breve{a}^j(i)\right) = \lambda \bar{V}_i\left(c_i(\ell_i)\right) + (1-\lambda) \bar{V}_i\left(c_i(\ell_i')\right) \\
&= \lambda \bar{V}_i(\ell_i) + (1-\lambda) \bar{V}_i(\ell_i')
\end{aligned}$$

Summarizing,

$$\bar{V}_i(\lambda \ell_i + (1-\lambda) \ell_i') = \lambda \bar{V}_i(\ell_i) + (1-\lambda) \bar{V}_i(\ell_i')$$

and the affinity of $\bar{V}_i$ is established.

Hence,

$$\bar{V}(\gamma, \ell) = \sum_{i=1}^{n} \gamma_i U_i(\ell_i)$$

as required. ■

6.3 Preference for identity lotteries vs source indifference

Lemma 3 If the impartial observer preferences satisfy *strong acceptance*, *convexity* and *preference for identity lotteries* then they satisfy *source indifference*.

Proof Consider, without loss of generality, a society $I = \{1, ..., n\}$, a set of available actions $A = \{a^1, ..., a^n\}$ and assume that (again, without loss of generality) $V(i, a^1) = V(1, a^i) := v_i$, for all i. Without loss of generality we can assume that all v_i are pairwise different and that $v_i > v_{i+1}$ for all $i < n$. For $i, j \in \{1, ..., n\}$, let $x_{ij} \in \mathcal{X}$ be defined by $V_i(x_{ij}) = v_{\pi_j(i)}$, where π_j is a permutation on $\{1, ..., n\}$ (as defined in the proof of the theorem), and note that, by the monotonicity of each V_i with respect to the outcomes of $\mathcal{X}$, $V_1(x_{11}) > V_1(x_{12}) > \cdots > V_1(x_{1n})$, $V_2(x_{2n}) > V_2(x_{21}) > V_2(x_{22}) > \cdots > V_2\left(x_{2(n-1)}\right)$,...,$V_n(x_{n2}) > V_n(x_{n3}) > \cdots > V_n(x_{nn}) > V_n(x_{n1})$. Consider a new set of actions $\bar{A} = \{\bar{a}^1, ..., \bar{a}^n\}$ satisfying $\bar{a}^j(i) = x_{ij}$. By construction,

$$V\left(i, \bar{a}^1\right) = V_i(x_{i1}) = v_{\pi_1(i)} = v_i = V\left(i, a^1\right)$$

and

$$V\left(1,\bar{a}^i\right) = V_1\left(x_{1i}\right) = v_{\pi_i(1)} = v_i = V\left(1,a^i\right)$$

which implies that, by strong acceptance, $V\left(\gamma^e,a^1\right) = W\left(\gamma^e,\left(v_1,...,v_n\right)\right) = V\left(\gamma^e,\bar{a}^1\right)$ and $V\left(1,\ell^e\right)$, given A, is equal to $V\left(1,\ell^e\right)$, given $\bar{A}$. Hence it is sufficient to restrict attention to $\bar{A}$ and to show that $V\left(\gamma^e,\bar{a}^1\right) = V\left(1,\ell^e\right)$ (given $\bar{A}$). For this note that: (i) since $V\left(\gamma^e,\bar{a}^i\right) = W\left(\gamma^e,\left(v_1,...,v_n\right)\right)$ for all i, we have $V\left(\gamma^e,\bar{a}^i\right) = V\left(\gamma^e,\bar{a}^j\right)$, for all i,j; (ii) by construction, for every $k\in\{1,...,n\}$, $V\left(i,\bar{a}^k\right) = V\left(k,\bar{a}^i\right)$, for all i; (iii) $V\left(\gamma^e,\ell^e\right)\in\left[\min_i V\left(i,\ell^e\right),\max_i V\left(i,\ell^e\right)\right]$ and hence, if $V\left(\gamma^e,\ell^e\right)=\max_i V\left(i,\ell^e\right)$ then $V\left(\gamma^e,\ell^e\right) = V\left(i,\ell^e\right) = V\left(j,\ell^e\right)$, for all i,j; and (iv) individual i strictly prefers action $\bar{a}^{n+2-i}$ (where $\bar{a}^{n+2-1} = \bar{a}^{n+1} := \bar{a}^1$) over all other actions and, by the monotonicity of V_i with respect to first-order stochastic-dominance, he strictly prefers action $\bar{a}^i$ over all mixtures of the other actions. Therefore,

$$V\left(\gamma^e,\bar{a}^1\right) = \max_k V\left(\gamma^e,\bar{a}^k\right) \geqslant \max_k V\left(k,\ell^e\right) \geqslant V\left(\gamma^e,\ell^e\right) \geqslant V\left(\gamma^e,\bar{a}^1\right)$$

where the equality follows from (i), the first inequality follows from (ii) and from preference for identity lotteries, the second inequality follows from the first part of (iii) and the last inequality follows from (iv) by repeated application of convexity (note that $\ell^e = \frac{1}{n}\sum_j \bar{a}^j$).

Since the first and the last elements are identical, $\max_k V\left(k,\ell^e\right) = V\left(\gamma^e,\ell^e\right)$ which, by the second part of (iii), implies that $V\left(1,\ell^e\right) = \max_k V\left(k,\ell^e\right)$ and, therefore, $V\left(1,\ell^e\right) = V\left(\gamma^e,\bar{a}^1\right)$. Hence the impartial observer is indifferent between identity and action lotteries. ■

6.4 GKPS's (2010) source indifference implies ours

Lemma 4 Assume (as in GKPS 2010) that the impartial observer satisfies the following property:

$\forall\gamma,\gamma'\in\triangle(\mathcal{I})$, $\forall\ell,\ell'\in\triangle(\mathcal{A})$ and $\forall\beta\in(0,1)$,

$$\left(\gamma,\ell'\right)\sim\left(\gamma',\ell\right)\Rightarrow\left(\beta\gamma+\left(1-\beta\right)\gamma',\ell\right)\sim\left(\gamma,\beta\ell+\left(1-\beta\right)\ell'\right)$$

Then the impartial observer exhibits *source indifference.*

Proof. The proof is by induction. Without loss of generality, consider a society $I = \{1, ..., n\}$, the set of available actions $A = \{a^1, ..., a^n\}$ and assume that $(1, a^i) \sim (i, a^1)$, for all i.

First let $n = 2$. By the GKPS. condition, $(1, a^2) \sim (2, a^1)$ implies

$$(\frac{1}{2}1 + \frac{1}{2}2, a^1) \sim (1, \frac{1}{2}a^1 + \frac{1}{2}a^2)$$

as required.

Next assume it holds for $n - 1$ and consider n. Assume, without loss of generality, that the acts of A satisfy $(i, a^j) \sim (i + 1, a^{j-1})$ for all $i \in \{1, ..., n - 1\}$, $j \in \{2, ..., n\}$. Consider the society $I^{\setminus 1} = \{2, ..., n\}$ and the set of actions $A^{\setminus n} = \{a^1, ..., a^{n-1}\}$. By construction, $(2, a^i) \sim (i + 1, a^1)$ for all $i = 1, ..., n - 1$ and hence, by the induction hypothesis, $(\frac{1}{n-1}\sum_{i=2}^{n} i, a^1) \sim (2, \frac{1}{n-1}\sum_{i=1}^{n-1} a^i)$. Next apply the same argument to $I^{\setminus n} = \{1, ..., n - 1\}$ and $A^{\setminus n} = \{a^1, ..., a^{n-1}\}$, where $(2, a^i) \sim (i, a^2)$ for all i, to get $(2, \frac{1}{n-1}\sum_{i=1}^{n-1} a^i) \sim (\frac{1}{n-1}\sum_{i=1}^{n-1} i, a^2)$. Finally, apply it to $I^{\setminus n} = \{1, ..., n - 1\}$ and $A^{\setminus 1} = \{a^2, ..., a^n\}$, where $(1, a^{i+1}) \sim (i, a^2)$ for all i, to get $(\frac{1}{n-1}\sum_{i=1}^{n-1} i, a^2) \sim (1, \frac{1}{n-1}\sum_{i=2}^{n} a^i)$. By transitivity,

$$(\frac{1}{n-1}\sum_{i=2}^{n} i, a^1) \sim (1, \frac{1}{n-1}\sum_{i=2}^{n} a^i)$$

To conclude, mix both sides of the last indifference with $(1, a^1)$ and, by the GKPS. condition, obtain $(\alpha^e, a^1) \sim (1, \ell^e)$ for $I = \{1, ..., n\}$ and $A = \{a^1, ..., a^n\}$, as required. ■

References

[1] Blackorby, C., D. Donaldson and P. Mongin. 2004. Social aggregation without the expected utility hypothesis. Discussion paper 2004-020, Ecole Polytechnique.

[2] Chew, S.H. 1983. A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. *Econometrica*, 51, 1065-1092.

[3] Chew, S.H. 1989. Axiomatic utility theories with the betweenness property. *Annals of operations Research*, 19, 273-298.

[4] Chew, S.H., E. Karni and Z. Safra. 1987. Risk aversion in the theory of expected utility with rank dependent probabilities. *Journal of Economic Theory*, 42, 370-381.

[5] Dekel, E. 1986. An axiomatic characterization of preferences under uncertainty: weakening the independence axiom. *Journal of Economic Theory*, 40, 304-318.

[6] Dekel, E., Z. Safra and U. Segal. 1991. Existence and dynamic consistency of Nash equilibrium with non-expected utility preferences. *Journal Economic Theory*, 55, 229-246.

[7] Dhillon, A. and J.F. Mertens. 1999. Relative utilitarianism. *Econometrica*, 67(3), 471-498.

[8] Diamond, P. A. 1967. Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: Comment. *The Journal of Political Economy*, 75(5), 765-766.

[9] Elster, J. 1989. *Solomonic Judgements: Studies in the Limitation of Rationality*. Cambridge University Press.

[10] Epstein, L. and U. Segal. 1992. Quadratic social welfare functions. *Journal of Political Economy*, 100(4), 691-712.

[11] Fleurbaey, M. and P. Mongin. 2016. The utilitarian relevance of the aggregation theorem. *American Economic Journal: Microeconomics*, 8(3), 289-306.

[12] Gilboa, I., D. Samet and D. Schmeidler. 2004. Utilitarian aggregation of beliefs and tastes. *Journal of Political Economy*, 112(4), 932-938.

[13] Gorman, W.M. 1968. The structure of utility functions. *Review of Economic Studies*, 35, 367-390.

[14] Grant, S., A. Kajii, B. Polak and Z. Safra. 2010. Generalized utilitarianism and Harsanyi's impartial observer theorem. *Econometrica*, 78(6), 1939-1971.

[15] Grant, S., A. Kajii, B. Polak and Z. Safra. 2012. A generalized representation theorem for Harsanyi's ('impartial') observer. *Social Choice and Welfare*, 39, 833–846.

[16] Harsanyi, J.C. 1953. Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy*, 61, 434-435.

[17] Harsanyi, J.C. 1955. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, 63, 309-321.

[18] Harsanyi, J.C. 1975. Nonlinear social welfare functions. *Theory and Decision*, 6, 311-332.

[19] Harsanyi, J.C. 1977. *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge University Press.

[20] Herstein, I. N. and J. Milnor. 1953. An axiomatic approach to measurable utility. *Econometrica*, 21(2), 291-297.

[21] Karni, E. 1998. Impartiality: definition and representation. *Econometrica*, 66(6), 1405-1415.

[22] Karni, E and Z. Safra. 2002. Individual sense of justice: a utility representation. *Econometrica*, 70, 263–284.

[23] Mongin, P. and M. Pivato. 2015. Ranking multidimensional alternatives and uncertain prospects. *Journal of Economic Theory*, 157, 146-171.

[24] Quiggin J. 1982. A theory of anticipated utility. *Journal of Economic Behavior and Organization*, 3, 323-343.

[25] Quiggin J. 1993. *Generalized Expected Utility Theory: The Rank-Dependent Model*. Kluwer Academic Publishers.

[26] Segal, U. 2000. Let's agree that all dictatorships are equally bad. *Journal of Political Economy*, 108(3), 569-589.

[27] Sen, A.K. 1976. Welfare inequalities and Rawlsian axiomatics. *Theory and Decision*, 7, 243-262.

[28] Sen, A.K. 1977. Non-linear social welfare functions: A reply to Professor Harsanyi. In *Foundational Problems in the Social Sciences*, R. Butts and J. Hintikka (eds.). Reidel Publishing Company, 297-302.

[29] Weymark, J.A. 1981. Generalized Gini inequality indices. *Mathematical Social Sciences*, 1(4), 409-430.

[30] Weymark, J.A. 1991. A reconsideration of the Harsanyi-Sen debate on utilitarianism. In *Interpersonal Comparisons of Well-being*, J. Elster and J.E. Roemer (eds). Cambridge University Press, 255-320.

[31] Zhou, L. 1997. Harsanyi's utilitarianism theorems: general societies. *Journal of Economic Theory*, 72(1), 198-207.